

## A Threat in the Air

### How Stereotypes Shape Intellectual Identity and Performance

*Claude M. Steele*

From an observer's standpoint, the situations of a boy and a girl in a math classroom or of a Black student and a White student in any classroom are essentially the same. The teacher is the same; the textbooks are the same; and in better classrooms, these students are treated the same. Is it possible, then, that they could still experience the classroom differently, so differently in fact as to significantly affect their performance and achievement there? This is the central question of this article, and in seeking an answer, it has both a practical and a theoretical focus. The practical focus is on the perhaps obvious need to better understand the processes that can hamper a group's school performance and on what can be done to improve that performance. The theoretical focus is on how societal stereotypes about groups can influence the intellectual functioning and identity development of individual group members. To show the generality of these processes and their relevance to important outcomes, this theory is applied to two groups: African Americans, who must contend with negative stereotypes about their abilities in many scholastic domains, and women, who must do so primarily in math and the physical sciences. In trying to understand the schooling outcomes of these two groups, the theory has a distinct perspective, that of viewing people, in Sartre's (1946/1965) words, as "first of all beings in a situation" such that if one wants to understand them, one "must inquire first into the situation surrounding [them]" (p. 60).

The theory begins with an assumption: that to sustain school success one must be identified with school achievement in the sense of its being a part of one's self-definition, a personal identity to which one is self-evaluatively accountable. This accountability—that good self-feelings depend in some part on good achievement—translates into sustained achievement motivation. For such an identification to form, this reasoning continues, one must perceive good prospects in the domain, that is, that one has the interests, skills, resources, and opportunities to prosper there, as well as that one belongs there, in the sense of being accepted and valued in the domain. If this relationship to schooling does not form or gets broken, achievement may suffer. Thus, in trying to understand what imperils achievement among women and African Americans, this logic points to a basic question: What in the experience of these groups might frustrate their identification with all or certain aspects of school achievement?

One must surely turn first to social structure: limits on educational access that

have been imposed on these groups by socioeconomic disadvantage, segregating social practices, and restrictive cultural orientations, limits of both historical and ongoing effect. By diminishing one's educational prospects, these limitations (e.g., inadequate resources, few role models, preparational disadvantages) should make it more difficult to identify with academic domains. To continue in math, for example, a woman might have to buck the low expectations of teachers, family, and societal gender roles in which math is seen as unfeminine as well as anticipate spending her entire professional life in a male-dominated world. These realities, imposed on her by societal structure, could so reduce her sense of good prospects in math as to make identifying with it difficult.

But this article focuses on a further barrier, one that has its effect on the already identified, those members of these groups who, having survived structural obstacles, have achieved identification with the domain (of the present groups, school-identified African Americans and math-identified women). It is the social-psychological threat that arises when one is in a situation or doing something for which a negative stereotype about one's group applies. This predicament threatens one with being negatively stereotyped, with being judged or treated stereotypically, or with the prospect of conforming to the stereotype. Called *stereotype threat*, it is a situational threat—a threat in the air—that, in general form, can affect the members of any group about whom a negative stereotype exists (e.g., skateboarders, older adults, White men, gang members). Where bad stereotypes about these groups apply, members of these groups can fear being reduced to that stereotype. And for those who identify with the domain to which the stereotype is relevant, this predicament can be self-threatening.

Negative stereotypes about women and African Americans bear on important academic abilities. Thus, for members of these groups who are identified with domains in which these stereotypes apply, the threat of these stereotypes can be sharply felt and, in several ways, hampers their achievement.

First, if the threat is experienced in the midst of a domain performance—classroom presentation or test-taking, for example—the emotional reaction it causes could directly interfere with performance. My colleagues and I (Spencer, Steele, & Quinn, 1997; C. M. Steele & Aronson, 1995) have tested this possibility with women taking standardized math tests and African Americans taking standardized verbal tests. Second, when this threat becomes chronic in a situation, as for the woman who spends considerable time in a competitive, male-oriented math environment, it can pressure *disidentification*, a reconceptualization of the self and of one's values so as to remove the domain as a self-identity, as a basis of self-evaluation. Disidentification offers the retreat of not caring about the domain in relation to the self. But as it protects in this way, it can undermine sustained motivation in the domain, an adaptation that can be costly when the domain is as important as schooling.

Stereotype threat is especially frustrating because, at each level of schooling, it affects the vanguard of these groups, those with the skills and self-confidence to have identified with the domain. Ironically, their susceptibility to this threat derives not from internal doubts about their ability (e.g., their internalization of the stereotype) but from their identification with the domain and the resulting concern they have about being stereotyped in it. (This argument has the hopeful implication that to

improve the domain performance of these students, one should focus on the feasible task of lifting this situational threat rather than on altering their internal psychology.) Yet, as schooling progresses and the obstacles of structure and stereotype threat take their cumulative toll, more of this vanguard will likely be pressured into the ranks of the unidentified. These students, by not caring about the domain vis-à-vis the self, are likely to underperform in it regardless of whether they are stereotype threatened there. Thus, although the identified among these groups are likely to underperform only under stereotype threat, the unidentified (casualties of sociocultural disadvantage or prior internalization of stereotype threat) are likely to underperform and not persist in the domain even when stereotype threat has been removed.

In these ways, then, the present analysis sees social structure and stereotypes as shaping the academic identities and performance outcomes of large segments of society. But first, for the two groups under consideration, what are these outcomes?

As is much discussed, these outcomes are in a crisis state for African Americans. Although Black students begin school with standardized test scores that are not too far behind those of their White counterparts, almost immediately a gap begins to appear (e.g., Alexander & Entwistle, 1988; Burton & Jones, 1982; Coleman et al., 1966) that, by the sixth grade in most school districts, is two full grade levels (Gerard, 1983). There have been encouraging increases in the number of African Americans completing high school or its equivalence in recent years: 77% for Black students versus 83% for White students (American Council on Education, 1995–1996). And there have been modest advances in the number of African American high school graduates enrolling in college, although these have not been as substantial as in other groups (American Council on Education, 1995–1996). Perhaps most discouraging has been the high dropout rate for African American college students: Those who do not finish college within six years is 62%, compared with a national dropout rate of 41% (American Council on Education, 1995–1996). And there is evidence of lower grade performance among those who do graduate of, on average, two thirds of a letter grade lower than those of other graduating students (Nettles, 1988). On predominantly White campuses, Black students are also underrepresented in math and the natural sciences. Although historically Black colleges and universities now enroll only 17% of the nation's Black college students, they produce 42% of all Black BS degrees in natural science (Culotta & Gibbons, 1992). At the graduate level, although Black women have recently shown modest gains in PhDs received, the number awarded to Black men has declined over the past decade more than for any other subgroup in society (American Council on Education, 1995–1996).

Women clearly thrive in many areas of schooling. But in math, engineering, and the physical sciences, they often endure lesser outcomes than men. In a meta-analysis involving over 3 million participants, Hyde, Fennema, and Lamon (1990), for example, found that through elementary and middle school, there are virtually no differences between boys and girls in performance on standardized math tests but that a trend toward men doing better steadily increases from high school ( $SD = .29$ ) through college ( $SD = .41$ ) and into adulthood ( $SD = .59$ ). And, as their college careers begin, women leave these fields at a rate two and a half times that of men (Hewitt & Seymour, 1991). Although White women constitute 43% of the U.S. population, they earn only 22% of the BS degrees and 13% of the PhDs and occupy only

10% of the jobs in physical science, math, and engineering, where they earn only 75% of the salary paid to men (Hewitt & Seymour, 1991).

These inequities have compelled explanations ranging from the sociocultural to the genetic. In the case of African Americans, for example, past and ongoing socioeconomic disadvantage, cultural orientations (e.g., Ogbu, 1986), and genetic differences (e.g., Herrnstein & Murray, 1994; Jensen, 1969) have all been proposed as factors that, through singular and accumulated effect, could undermine their performance. In the case of women's performance in math and the physical sciences, there are parallel arguments: structural and cultural gender role constraints that shunt women away from these areas; culturally rooted expectations (e.g., Eccles, 1987; Eccles-Parsons et al., 1983); and, again, genetic limitations (Benbow & Stanley, 1980, 1983). But, like crumbs along the forest floor, several findings lead away from these analyses as fully sufficient.

For one thing, minority student achievement gaps persist even in the middle and upper socioeconomic classes. Using data from the Coleman report (Coleman et al., 1966) and a more recent College Board study of Scholastic Assessment Test (SAT) scores, Miller (1995, 1996) found that the gaps in academic performance (grades as well as standardized test scores) between Whites and non-Asian minorities (e.g., African Americans, Hispanics, and Native Americans) were as large, or larger, in the upper and middle classes (as measured by parental education and occupation) than in the lower classes. Group differences in socioeconomic status (SES), then, cannot fully explain group differences in academic performance.

Another point is that these differences are not even fully explained by group differences in skills. This is shown in the well-known *overprediction* or *underperformance* phenomenon of the test bias literature. Overprediction occurs when, at each level of performance on a test of preparation for some level of schooling (e.g., the SAT), students from one group wind up achieving less—getting lower college grades, for example—than other students with the same beginning scores. In this sense, the test scores of the low-performing group overpredict how well they will actually achieve, or, stated another way, the low-performing group underperforms in relation to the test's prediction. But the point here is that because the students at each test-score level have comparable initial skills, the lower eventual performance of one group must be due to something other than skill deficits they brought with them.

In the case of African Americans, overprediction across the academic spectrum has been so reliably observed as to be almost a lawful phenomenon in American society (e.g., Jensen, 1980; Vars & Bowen, 1997). Perhaps the most extensive single demonstration of it comes from a recent Educational Testing Service study (Ramist, Lewis, & McCamley-Jenkins, 1994) that examined the predictiveness of the SAT on 38 representative college and university campuses. As is typically the case, the study found that the predictive validity to the SAT—its correlation with subsequent grades—was as good for African American, Hispanic, and Native American students as for White and Asian students. But for the three non-Asian minority groups, there was sizable overprediction (underperformance) in virtually all academic areas. That is, at each level of preparation as measured by the SAT, something further depressed the grades of these groups once they arrived on campus.

As important, the same study found evidence of SAT overprediction for female



students (i.e., women performing less well than men at comparable SAT levels) in technical and physical science courses such as engineering, economics, and computer science but not in non-technical areas such as English. It is interesting though that women in this study were not overpredicted in math *per se*, a seeming exception to this pattern. The overprediction of women's college math performance has generally been unreliable, with some studies showing it (e.g., Benbow & Arjmand, 1990; Levin & Wyckoff, 1988; Lovely, 1987; Ware, Steckler, & Leserman, 1985) and others not (e.g., Adelman, 1991; DeBoer, 1984; Ware & Dill, 1986). However, a recent study (Strenta, Elliott, Adair, Scott, & Matier, 1993) involving over 5,000 students at four prestigious northeastern colleges identified a pattern of effects that suggests why these different results occur: Underperformance reliably occurred among women who were talented in math and science and who, perhaps for that reason, took courses in these areas that were intended for majors, whereas it did not occur among women with less math and science preparation who took courses in these areas intended for nonmajors. Thus, women may be reliably overpredicted in math and the physical sciences, just as Black students are more generally, but only when the curriculum is more advanced and only among women who are more identified with the domain. Among this vanguard, though, something other than skill deficits depresses their performance. What are these further processes?

### Social and Stereotype Structure as Obstacles to Achievement Identification

The proposed answer is that at least one of these processes is a set of social psychological phenomena that obstructs these groups' identification with domains of schooling.<sup>1</sup> I turn first to school identification.

#### Academic Identification

As noted, this analysis assumes that sustained school achievement depends, most centrally, on identifying with school, that is, forming a relationship between oneself and the domains of schooling such that one's self-regard significantly depends on achievement in those domains. Extrinsic rewards such as better career outcomes, personal security, parental exhortation, and so on, can also motivate school achievement. But it is presumed that sustaining motivation through the ebb and flow of these other rewards requires school identification. How, then, is this identification formed?

Not a great deal is known about the process. But several models (e.g., Schlenker & Weigold, 1989; C. M. Steele, 1988; Tesser, 1988) share an implicit reasoning, the first assumption of which is that people need positive self-regard, a self-perception of "adaptive and moral adequacy" (C. M. Steele, 1988, p. 289). Then, the argument goes, identification with a given domain of life depends, in large part, on the self-evaluative prospects it offers. James (1890/1950) described the development of the self as a process of picking from the many, often incompatible, possible selves, those "on which to stake one's salvation" (p. 310). This choice and the assessment of prospects

that goes into it are, of course, multifaceted: Are the rewards of the domain attractive or important? Is an adequate opportunity structure available? Do I have the requisite skills, talents, and interests? Have others like me succeeded in the domain? Will I be seen as belonging in the domain? Will I be prejudiced against in the domain? Can I envision wanting what this domain has to offer? and so on. Some of these assessments undergird a sense of efficacy in the domain (e.g., Bandura, 1977, 1986). Others have to do with the rewards, importance, and attractiveness of the domain itself. And still others have to do with the feasibility and receptiveness of the domain. The point here is that students tacitly assess their prospects in school and its subdomains, and, roughly speaking, their identifications follow these assessments: increasing when they are favorable and decreasing when they are unfavorable. As for the two groups under consideration, then, this analysis suggests that something systematically downgrades their assessments of, and thus their identification with, critical domains of schooling.

### Threats to Academic Identification

*Structural and cultural threats.* Both groups have endured and continue to endure sociocultural influences that could have such effects. Among the most replicable facts in the schooling literature is that SES is strongly related to school success and cognitive performance (e.g., Coleman et al., 1966; Miller, 1996). And because African Americans have long been disproportionately represented in lower socioeconomic classes, this factor surely contributes to their achievement patterns in school, both through the material limitations associated with lower SES (poor schools, lack of resources for school persistence, etc.) and through the ability of these limitations, by downgrading school-related prospects, to undermine identification with school. And beyond socioeconomic structure, there are cultural patterns within these groups or in the relation between these groups and the larger society that may also frustrate their identification with school or some part of it, for example, Ogbu's (1986) notion of a lower-class Black culture that is "oppositional" to school achievement or traditional feminine gender roles that eschew math-related fields (e.g., Eccles-Parsons et al., 1983; Linn, 1994).

*Stereotype threat.* Beyond these threats, waiting for those in these groups who have identified with school, is yet another threat to their identification, more subtle perhaps but nonetheless profound: that of stereotype threat. I define it as follows: the event of a negative stereotype about a group to which one belongs becoming self-relevant, usually as a plausible interpretation for something one is doing, for an experience one is having, or for a situation one is in, that has relevance to one's self-definition. It happens when one is in the *field* of the stereotype, what Cross (1991) called a "spotlight anxiety" (p. 195), such that one can be judged or treated in terms of a racial stereotype. Many analysts have referred to this predicament and the pressure it causes (e.g., Allport, 1954; Carter, 1991; Cose, 1993; Goffman, 1963; Howard & Hammond, 1985; E. E. Jones et al., 1984; Sartre, 1946/1965; C. M. Steele, 1975; C. M. Steele & Aronson, 1995; S. Steele, 1990). The present definition stresses that for a negative stereotype to be threatening, it must be self-relevant. Then, the situational contingency it establishes—the possibility of conforming to the stereotype or of

being treated and judged in terms of it—becomes self-threatening. It means that one could be limited or diminished in a domain that is self-definitional. For students from groups in which abilities are negatively stereotyped in all or some school domains and yet who remain identified with those domains, this threat may be keenly felt, felt enough, I argue, to become a further barrier to their identification with the domain.

There is, however, a more standard explanation of how negative stereotypes affect their targets. Beginning with Freud (as cited in Brill, 1938) in psychology and Cooley (1956) and Mead (1934) in sociology, treatises on the experience of oppression have depicted a fairly standard sequence of events: Through long exposure to negative stereotypes about their group, members of prejudiced-against groups often internalize the stereotypes, and the resulting sense of inadequacy becomes part of their personality (e.g., Allport, 1954; Bettelheim, 1943; Clark, 1965; Grier & Coobs, 1968; Erikson, 1956; Fanon, 1952/1967; Kardiner & Ovesey, 1951; Lewin, 1941).

In recent years, the tone of this argument has constructively lightened, replacing the notion of a broad self-hatred with the idea of an inferiority anxiety or low expectations and suggesting how situational factors contribute to this experience. S. Steele's (1990) essays on *racial vulnerability* (i.e., a vulnerability of both Blacks and Whites that stems, in part, from the situational pressures of reputations about their groups) offered an example. This work depicts the workings of this anxiety among African Americans in an interconnected set of ideas: *integration shock* that, like Goffman (1963), points to settings that integrate Blacks and Whites as particularly anxiety arousing; *objective correlatives* or race-related situational cues that can trigger this anxiety; and the inherent sense of risk, stemming from an internalized *inferiority anxiety* and from a *myth of inferiority* pervading integrated settings, of being judged inferior or of confirming one's own feared inferiority. Howard and Hammond (1985) earlier made this argument specifically in relation to the school achievement of Black students. They argued that once "rumors of inferiority" (stereotypes; p. 18) about Black students' abilities pervade the environment—through, for example, national debates over the genetic basis of racial differences in IQ—they can intimidate Black students; become internalized by them; and, in turn, lead to a low sense of self-efficacy, demotivation, and underperformance in school. Analogous arguments have been applied to women interested in math-related areas (cf. Eccles-Parsons et al., 1983).

These models recognize the situational influence of negative stereotypes (e.g., Allport, 1954; Howard & Hammond, 1985; S. Steele, 1990) but most often describe it as a process in which the stereotype, or more precisely the possibility of being stereotyped, triggers an internalized inferiority doubt or low expectancy. And because this anxiety is born of a socialization presumed to influence all members of the stereotyped group, virtually all members of the group are presumed to have this anxiety, to one degree or another.

Stereotype threat, in contrast, refers to the strictly situational threat of negative stereotypes, the threat that does not depend on cuing an internalized anxiety or expectancy. It is cued by the mere recognition that a negative group stereotype could apply to oneself in a given situation. How threatening this recognition becomes depends on the person's identification with the stereotype-relevant domain. For the

domain identified, the situational relevance of the stereotype is threatening because it threatens diminishment in a domain that is self-definitional. For the less domain identified, this recognition is less threatening or not threatening at all, because it threatens something that is less self-definitional.

Stereotype threat, then, as a situational pressure "in the air" so to speak, affects only a subportion of the stereotyped group and, in the area of schooling, probably affects confident students more than unconfident ones. Recall that to be identified with schooling in general, or math in particular, one must have confidence in one's domain-related abilities, enough to perceive good prospects in the domain. This means that stereotype threat should have its greatest effect on the better, more confident students in stereotyped groups, those who have not internalized the group stereotype to the point of doubting their own ability and have thus remained identified with the domain—those who are in the academic vanguard of their group.<sup>2</sup>

Several general features of stereotype threat follow:

1. Stereotype threat is a general threat not tied to the psychology of particular stigmatized groups. It affects the members of any group about whom there exists some generally known negative stereotype (e.g., a grandfather who fears that any faltering of memory will confirm or expose him to stereotypes about the aged). Stereotype threat can be thought of as a subtype of the threat posed by negative reputations in general.

2. That which turns stereotype threat on and off, the controlling "mechanism" so to speak, is a particular concurrence: whether a negative stereotype about one's group becomes relevant to interpreting oneself or one's behavior in an identified-with setting. When such a setting integrates stereotyped and nonstereotyped people, it may make the stereotype, as a dimension of difference, more salient and thus more strongly felt (e.g., Frible, Blackstone, & Sherbaum, 1990; Goffman, 1963; Kleck & Strenta, 1980; Sartre, 1946/1965; S. Steele, 1990). But such integration is neither necessary nor sufficient for this threat to occur. It can occur even when the person is alone, as for a woman taking an important math test alone in a cubicle but under the threat of confirming a stereotyped limitation of ability. And, in integrated settings, it need not occur. Reducing the interpretive relevance of a stereotype in the setting, say in a classroom or on a standardized test, may reduce this threat and its detrimental effects even when the setting is integrated.<sup>3</sup>

3. This mechanism also explains the variabilities of stereotype threat: the fact that the type and degree of this threat vary from group to group and, for any group, across settings. For example, the type and degree of stereotype threat experienced by White men, Black people, and people who are overweight differ considerably, bearing on sensitivity and fairness in the first group, on school performance in the second, and on self-control in the third. Moreover, for any of these groups, this threat will vary across settings (e.g., Goffman, 1963; S. Steele, 1990). For example, women may reduce their stereotype threat substantially by moving across the hall from math to English class. The explanation of this model is straightforward: Different groups experience different forms and degrees of stereotype threat because the stereotypes about them differ in content, in scope, and in the situations to which they apply.

4. To experience stereotype threat, one need not believe the stereotype nor even be



worried that it is true of oneself. The well-known African American social psychologist James M. Jones (1997) wrote,

When I go to the ATM machine and a woman is making a transaction, I think about whether she will fear I may rob her. Since I have no such intention, how do I put her at ease? Maybe I can't . . . and maybe she has no such expectation. But it goes through my mind. (p. 262)

Jones felt stereotype threat in this situation even though he did not believe that the stereotype characterized him. Of course, this made it no less a life-shaping force. One's daily life can be filled with recurrent situations in which this threat pressures adaptive responses.

5. The effort to overcome stereotype threat by disproving the stereotype—for example, by outperforming it in the case of academic work—can be daunting. Because these stereotypes are widely disseminated throughout society, a personal exemption from them earned in one setting does not generalize to a new setting where either one's reputation is not known or where it has to be renegotiated against a new challenge. Thus, even when the stereotype can be disproven, the need to do so can seem Sisyphean, everlastingly recurrent. And in some critical situations, it may not be disprovable. The stereotypes considered in this work allege group-based limitations of ability that are often reinforced by the structural reality of increasingly small group representations at more advanced levels of the schooling domain. Thus, for group members working at these advanced levels, no amount of success up to that point can disprove the stereotype's relevance to their next, more advanced performance. For the advanced female math student who has been brilliant up to that point, any frustration she has at the frontier of her skills could confirm the gender-based limitation alleged in the stereotype, making this frontier, because she is so invested in it, a more threatening place than it is for the nonstereotyped. Thus, the work of dispelling stereotype threat through performance probably increases with the difficulty of work in the domain, and whatever exemption is gained has to be won at the next new proving ground.

### Empirical Support for a Theory of Stereotype Threat and Disidentification

In testing these ideas, the research of my colleagues and I has had two foci: The first is on intellectual performance in the domain in which negative group stereotypes apply. Here, the analysis has two testable implications. One is that for domain-identified students, stereotype threat may interfere with their domain-related intellectual performance. Analysts have long argued that behaving in a situation in which one is at risk of confirming a negative stereotype about one's group, or of being seen or treated stereotypically, causes emotional distress and pressure (e.g., Cross, 1991; Fanon, 1952/1967; Goffman, 1963; Howard & Hammond, 1985; Sartre, 1946/1965; C. M. Steele & Aronson, 1995; S. Steele, 1990). The argument here is that for those who identify with the domain enough to experience this threat, the pressure it causes may undermine their domain performance. Disruptive pressures such as evaluation apprehension, test anxiety, choking, and token status have long been shown to

disrupt performance through a variety of mediating mechanisms: interfering anxiety, reticence to respond, distracting thoughts, self-consciousness, and so on (Baumeister & Showers, 1984; Geen, 1991; Lord & Saenz, 1985; Sarason, 1980; Wine, 1971). The assumption of this model is that stereotype threat is another such interfering pressure. The other testable implication is that reducing this threat in the performance setting, by reducing its interfering pressure, should improve the performance of otherwise stereotype-threatened students.

The second research focus is the model's implication that stereotype threat, and the anticipation of having to contend with it unceasingly in school or some domain of schooling, should deter members of these groups from identifying with these domains, and, for group members already identified, it should pressure their disidentification.<sup>4</sup>

### Stereotype Threat and Intellectual Performance

Steven Spencer, Diane Quinn, and I (Spencer et al., 1997) first tested the effect of stereotype threat on intellectual performance by testing its effect on the standardized math test performance of women who were strong in math.

*The stereotype threat of women performing math.* At base, of course, the stereotype threat that women experience in math-performance settings derives from a negative stereotype about their math ability that is disseminated throughout society. But whether this threat impaired their performance, we reasoned, would depend on two things. First, the performance would have to be construed so that any faltering would imply the limitation of ability alleged in the stereotype. This means that the performance would have to be difficult enough so that faltering at it would imply having reached an ability limit but not so difficult as to be nondiagnostic of ability. And second, as has been much emphasized, the women in question would have to be identified with math, so that faltering and its stereotype-confirming implication would threaten something they care about, their belongingness and acceptance in a domain they identify with. Of course, men too (at least those of equal skill and identification with math) could be threatened in this situation; faltering would reflect on their ability too. But their faltering would not carry the extra threat of confirming a stereotyped limitation in math ability or of causing them to be seen that way. Thus, the threat that women experience, through the interfering pressure it causes, should worsen their performance in comparison to equally qualified men. Interestingly, though, these otherwise confident women should perform equally as well as equally qualified men when this situational threat is lessened.

To explore these questions, Spencer, Quinn, and I (Spencer et al., 1997) designed a basic research paradigm: We recruited female and male students, mostly college sophomores, who were both good at math and strongly identified with it in the sense of seeing themselves as strong math students and seeing math as important to their self-definition. We then gave them a very difficult math test one at a time. The items were taken from the advanced math General Records Examination (GRE) and we assumed would frustrate the skills of these students without totally exceeding them. As expected, and presumably reflecting the impairing effects of stereotype threat,

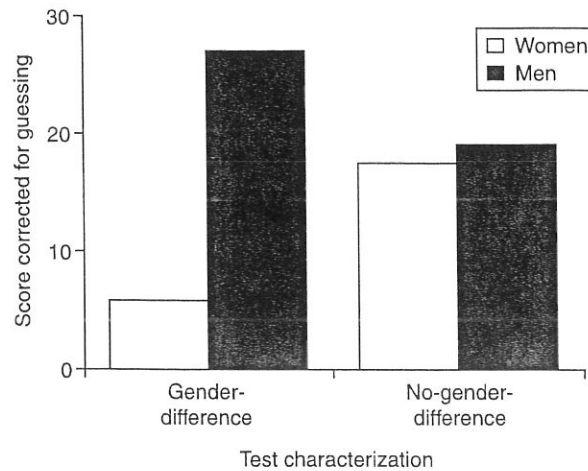
women significantly underperformed in relation to equally qualified men on this difficult math test. But more important, in another condition of this experiment in which the test was an advanced literature test rather than a math test and in which participants had been selected and matched for their strong literature skills and identification, women performed just as well as equally qualified men. This happened, we reasoned, because women are not stereotype threatened in this area.

A second experiment replicated women's underperformance on the difficult math test and showed that it did not happen when the test was easier, that is when the items, taken from the regular quantitative section of the GRE, were more within the skills of these strong math students. The lack of performance frustration on this easier test, presumably, reduced women's stereotype threat by making the stereotype less relevant as an interpretation of their performance.

*Stereotype threat versus genes.* So went our interpretation. But an alternative was possible: The biological limits of women's math ability do not emerge until the material tested is difficult. It is this very pattern of evidence that Benbow and Stanley (1980, 1983) used to suggest a genetic limitation in women's math ability. Thus, the first two experiments reproduced the gender effects on math performance reported in the literature: that women underperform primarily in math and mainly when the material is difficult. But they fall short of establishing our interpretation.

To do this, we would need to give women and men a difficult math test (one capable of producing women's underperformance) but then experimentally vary stereotype threat, that is, vary how much women were at risk of confirming the stereotype while taking the test. A third experiment did this by varying how the test (the same difficult one used in the earlier experiments) was represented. Participants were told either that the test generally showed gender differences, implying that the stereotype of women's limitations in math was relevant to interpreting their own frustration, or that it showed no gender differences, implying that the gender stereotype was not relevant to their performance and thus could not be confirmed by it on this particular test. The no-gender-differences representation did not challenge the validity of the stereotype; it simply eliminated the risk that the stereotype could be fulfilled on this test. In the gender-differences condition, we expected women (still stereotype threatened) to underperform in relation to equally qualified men, but in the no-gender-differences condition, we expected women (with stereotype threat reduced) to perform equal to such men. The genetic interpretation, of course, predicts that women will underperform on this difficult test regardless of how it is represented.

In dramatic support of our reasoning, women performed worse than men when they were told that the test produced gender differences, which replicated women's underperformance observed in the earlier experiments, but they performed equal to men when the test was represented as insensitive to gender differences, even though, of course, the same difficult "ability" test was used in both conditions (see Figure 11.1). Genetic limitation did not cap the performance of women in these experiments. A fourth experiment showed that reducing stereotype threat (through the no-gender-differences treatment) raised women's performance to that of equally qualified men, even when participants' specific performance expectancies were set low, that is, when participants were led to expect poor test performance. Also, a fifth



**Figure 11.1** Mean Performance on a Difficult Math Test as a Function of Gender and Test Characterization.

experiment (that again replicated the treatment effects of the third experiment) found that participants' posttreatment anxiety, not their expectancies or efficacy, predicted their performance. Thus, the disruptive effect of stereotype threat was mediated more by the self-evaluative anxiety it caused than by its lowering of performance expectations or self-efficacy.

*Internal or situational threat.* These findings make an important theoretical and practical point: The gender-differences conditions (including those in which the possibility of gender differences was left to inference rather than stated directly) did not impair women's performance by triggering doubts they had about their math ability. For one thing, these women had no special doubts of this sort; they were selected for being very good at math and for reporting high confidence in their ability. Nor was this doubt a factor in their test performance. Recall that the math test was represented as an ability test in all conditions of these experiments. This means that in the no-gender-differences conditions, women were still at risk of showing their own math ability to be weak—the same risk that men had in these conditions. Under this risk (when their own math ability was on the line), they performed just as well as men. Whatever performance-impairing anxiety they had, it was no greater than that of equally qualified men. Thus, the gender-differences conditions (the normal condition under which people take these tests) could not have impaired their performance by triggering some greater internalized anxiety that women have about their own math ability—an anxiety acquired, for example, through prior socialization. Rather, this condition had its effect through situational pressure. It set up an interpretive frame such that any performance frustration signaled the possible gender-based ability limitation alleged in the stereotype. For these women, this signal challenged their belongingness in a domain they cared about and, as a possibly newly met limit to their ability, could not be disproven by their prior achievements, thus its interfering threat.

*The stereotype threat of African Americans on standardized tests.* Joshua Aronson and I (C. M. Steele & Aronson, 1995) examined these processes among African



American students. In these studies, Black and White Stanford University students took a test composed of the most difficult items on the verbal GRE exam. Because the participants were students admitted to a highly selective university, we assumed that they were identified with the verbal skills represented on standardized tests. The first study varied whether or not the stereotype about Black persons' intellectual ability was relevant to their performance by varying whether the test was presented as *ability-diagnostic*, that is, as a test of intellectual ability, or as *ability-nondiagnostic*, that is, as a laboratory problem-solving task unrelated to ability and thus to the stereotype about ability. Analysis of covariance was used to remove the influence of participants' initial skills, measured by their verbal SAT scores, on their test performance. This done, the results showed strong evidence of stereotype threat: Black participants greatly underperformed White participants in the diagnostic condition but equaled them in the nondiagnostic condition (see Figure 11.2). A second experiment produced the same pattern of results with an even more slight manipulation of stereotype threat: whether or not participants recorded their race on a demographic questionnaire just before taking the test (described as nondiagnostic in all conditions). Salience of the racial stereotype alone was enough to depress the performance of identified Black students (see Figure 11.3).

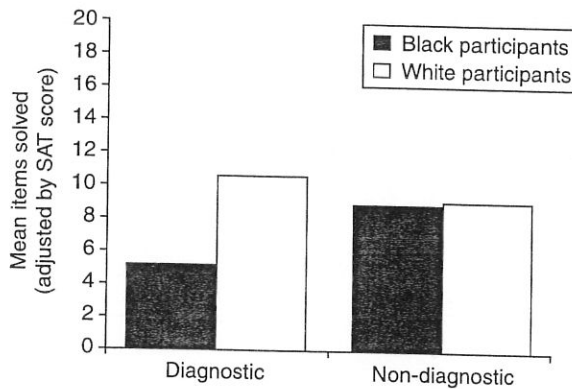


Figure 11.2 Mean Performance on a Difficult Verbal Test as a Function of Race and Test Characterization.

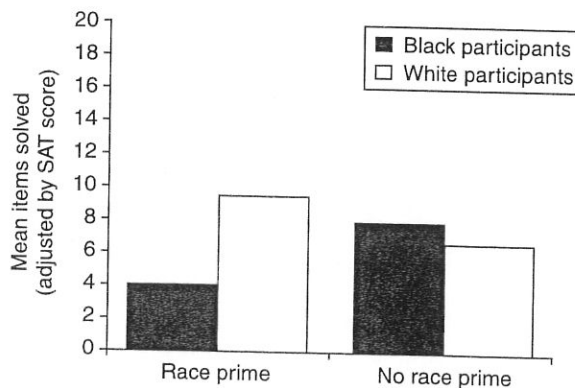


Figure 11.3 Mean Performance on a Difficult Verbal Test as a Function of Whether Race Was Primed.

*The cognitive mediation of stereotype threat.* Stereotype threat, then, can impair the standardized test performance of domain-identified students; this effect generalizes to several ability-stereotyped groups, and its mediation seems to involve anxiety more than expectancies. But do these manipulations cause a specific state of stereotype threat, that is, a sensed threat specifically about being stereotyped or fitting the stereotype? To address this question, Aronson and I (C. M. Steele & Aronson, 1995) tested two things: whether manipulating stereotype threat actually activates the racial stereotype in the thinking and information processing of stereotype-threatened test takers and whether it produces in them a specific motivation to avoid being seen stereotypically. Again, Black and White participants were run in either an ability-diagnostic or ability-nondiagnostic condition, except that just after the condition instructions and completion of the sample test items (so that participants could see how difficult the items were) and just before participants expected to take the test, they completed measures of stereotype activation and avoidance. The stereotype-activation measure asked them to complete 80 word fragments, 10 of which we knew from pretesting could be completed with, among other words, words symbolic of African American stereotypes (e.g., \_ \_ ce[race], la\_ \_ [lazy], or \_ \_ or [poor]) and 5 of which could be completed with, among other words, words signifying self-doubts (e.g., lo\_ \_ \_ [loser], du\_ \_ [dumb], or sha\_ \_ [shame]). The measure of participants' motivation to avoid being seen stereotypically simply asked them how much they preferred various types of music, activities, sports, and personality traits, some of which a pretest sample had rated as stereotypic of African Americans.<sup>5</sup>

If expecting to take a difficult ability-diagnostic test is enough to activate the racial stereotype in the thinking of Black participants and to motivate them to avoid being stereotyped, then these participants, more than those in the other conditions, should show more stereotype and self-doubt word completions and fewer preferences for things that are African American. This is precisely what happened. Black participants in the diagnostic condition completed more word fragments with stereotype- and self-doubt-related words and had fewer preferences for things related to African American experience (e.g., jazz, basketball, hip-hop) than Black participants in the nondiagnostic condition or White participants in either condition, all of whom were essentially the same (see Figure 11.4). Also, as a last item before participants expected to begin the test, they were given the option of recording their race, a measure we thought might further tap into an apprehension about being viewed stereotypically. Interestingly, then, all of the Black participants in the nondiagnostic condition and all of the White participants in both conditions listed their race, whereas only 25% of the Black participants in the diagnostic condition did so.

*Self-rejection or self-presentation?* A troubling implication of the earlier mentioned internalization models (e.g., Allport, 1954; Bettelheim, 1943; Clark, 1965; Grier & Coobs, 1968; Erikson, 1956; Fanon, 1952/1967; Kardiner & Ovesey, 1951) is that negative stereotypes about one's group eventually become internalized and cause rejection of one's own group, even of oneself—*self-hating* preferences. The famous finding of Clark and Clark (1939) that Black children preferred White dolls over Black dolls has been interpreted this way. The preferences of Black participants in the diagnostic condition fit this pattern; with negative stereotypes about their group cognitively activated, they valued things that were African American less than any

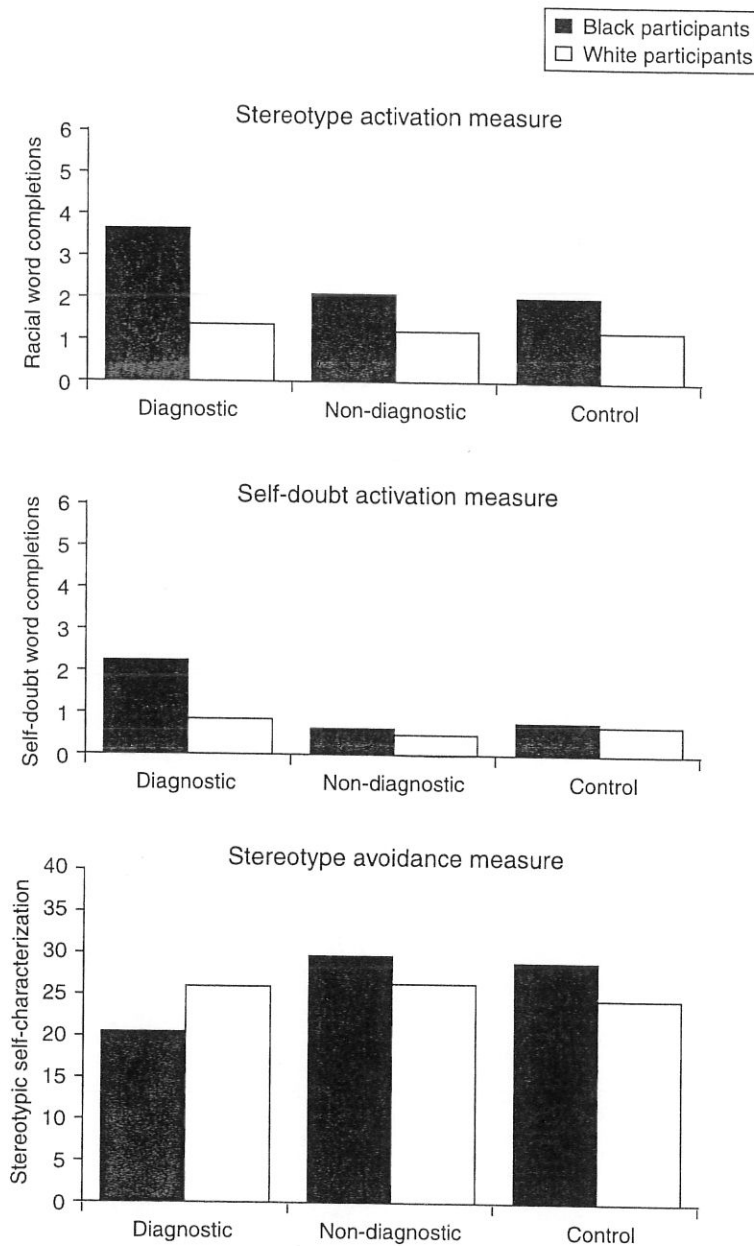


Figure 11.4 Indicators of Stereotype Threat.

other group. But the full set of results suggests a different interpretation. In those conditions in which Black participants did not have to worry about tripping a stereotypic perception of themselves, they valued things that were African American more strongly than did other participants. Thus, rather than reflecting self- or own-group rejection, their devaluing of things that were African American in the diagnostic condition was apparently a strategic self-presentation aimed at cracking the stereotypic lens through which they could be seen. So it could be, then, in the general

case, rather than reflecting real self-concepts, behavior that appears group rejecting or self-rejecting may reflect situation-bound, self-presentational strategies.

*Stereotype threat and domain identification.* Not being identified with a domain, our (C. M. Steele & Aronson, 1995) theory reasons, means that one's experience of stereotype threat in the domain is less self-threatening. Although we have yet to complete a satisfactory test of this prediction, partially completed experiments and pretests show that stereotype threat has very little, if any, effect on participants not identified with the domain of relevance. Most typically, these participants give up and underperform on the difficult test regardless of whether they are under stereotype threat. Although not yet constituting a complete test of this implication of the theory, these data do emphasize that the above results generalize only to domain-identified students.

*Stereotype threat and the interpretation of group differences in standardized test performance.* Inherent to the science of quantifying human intelligence is the unsavory possibility of ranking societal groups as to their aggregated intelligence. It is from this corner of psychology that the greatest controversy has arisen, a controversy that has lasted throughout this century and that is less about the fact of these group differences than about their interpretation (cf. Herrnstein & Murray, 1994; Kamin, 1974). To the set of possible causes for these group differences, our (C. M. Steele & Aronson, 1995) findings add a new one: the differential impact of stereotype threat on groups in the testing situation itself. Thus, stereotype threat may be a possible source of bias in standardized tests, a bias that arises not from item content but from group differences in the threat that societal stereotypes attach to test performance. Of course, not every member of an ability-stereotyped group is going to be affected by stereotype threat every time they take a test. As our research has shown, the experience of success as one takes the test can dispel the relevance of the stereotype. Nonetheless, among the most identified test takers in the stereotype-threatened group—those in its academic vanguard who have the greatest confidence and skills—this threat can substantially depress performance on more difficult parts of the exam. And this depression could contribute significantly to the group's underperformance in comparison with nonstereotype-threatened groups.<sup>6</sup>

### Reaction of Disidentification

Stereotype threat is assumed to have an abiding effect on school achievement—an effect beyond its impairment of immediate performance—by preventing or breaking a person's identification with school, in particular, those domains of schooling in which the stereotype applies. This reasoning has several implications for which empirical evidence can be brought to bear: the resilience of self-esteem to stigmatization; the relationship between stigmatized status and school achievement; and, among ability-stigmatized people, the relationship between their school performance and self-esteem.

*Self-esteem's resilience to stigmatization.* In a recent review, Crocker and Major (1989) were able to make a strong case for the lack of something that common sense suggests should exist: a negative effect of stigmatization on self-esteem. Following the



logic of the internalization models described above and viewing stigmatization as, among other things, an assault to self-esteem, one might expect that people who are stigmatized would have lower self-esteem than people who are not. Yet, as Crocker and Major reported, when the self-esteem of stigmatized groups (e.g., Blacks, Chicanos, the facially disfigured, obese people, etc.) is actually measured, one finds that their self-esteem is as high as that of the nonstigmatized.

Crocker and Major (1989) offered the intriguing argument that stigma itself offers esteem-protective strategies. For example, the stigmatized can blame their failures on the prejudice of out-group members, they can limit their self-evaluative social comparisons to the ingroup of other stigmatized people, and they can devalue the domains in which they feel devalued. Other models have also described esteem-saving adaptations to stigma. For example, models that assume internalization of stereotype-related anxieties often posit compensatory personality traits (e.g., grandiosity) that protect self-esteem but leave one poorly adapted to the mainstream (e.g., Allport, 1954; Clark, 1965; Grier & Coobs, 1968; Kardiner & Ovesey, 1951; S. Steele, 1990). In the present reasoning, stigmatization stems from stereotype threat in specific domains. Thus, it adds to the list of stigma adaptations the possibility of simple domain disidentification, the rescuing of self-esteem by rendering as self-evaluatively irrelevant the domain in which the stereotype applies. Herein may lie a significant source of the self-esteem resilience shown in stigmatized groups. This idea also implies that once domain disidentification is achieved, the pressure for adaptations of attribution and personality may be reduced.

*A universal connection between stigmatization and poor school achievement.* If disidentification with school, and the resulting underachievement, can be a reaction to ability-stigmatizing stereotypes in society, then it might be expected that ability stigmatization would be associated with poor school performance wherever it occurs in the world. Finding such a relationship would not definitively prove the present theory; the direction of causality could be quarreled with, as could the mediation of such a relationship. Still, it would be suggestive, and, in that respect, Ogbu (1986) reported an interesting fact: Among the caste-like minorities in industrial and non-industrial nations throughout the world (e.g., the Maoris of New Zealand, the Baraku of Japan, the Harijans of India, the Oriental Jews of Israel, and the West Indians of Great Britain), there exists the same 15-point IQ gap between them and the nonstigmatized members of their society as exists between Black and White Americans. These groups also suffer poorer school performance, higher dropout rates, and related behavior problems. Moreover, these gaps appear even when the stigmatized and nonstigmatized are of the same race, as in the case of the Baraku and other Japanese. What these groups share that is capable of explaining their deficits is a caste-like status that, through stereotypes in their societies, stigmatizes their intellectual abilities—sowing the seeds, I suggest, of their school disidentification.

*The disassociation of self-esteem and school achievement.* If the poor school achievement of ability-stigmatized groups is mediated by disidentification, then it might be expected that among the ability stigmatized, there would be a disassociation between school outcomes and overall self-esteem. Several kinds of evidence suggest this process among African Americans. First, there is the persistent finding that although Black students underperform in relation to White students on school

outcomes from grades to standardized tests (e.g., Demo & Parker, 1987; Simmons, Brown, Bush, & Blyth, 1978; C. M. Steele, 1992), their global self-esteem is as high or higher than that of White students (e.g., Porter & Washington, 1979; Rosenberg, 1979; Wylie, 1979). For both of these facts to be true, some portion of Black students must have acquired an imperviousness to poor school performance.

Several further studies suggest that this imperviousness is rooted in disidentification. In a study of desegregated schools in Champaign, Illinois, Hare and Costenell (1985) measured students' school achievement; overall self-esteem; and self-esteem in the specific domains of home life, school, and peer-group relations. Like others, they found that although Black students performed less well than White students, they still had comparable levels of overall self-esteem. Their domain-specific measures suggested why: Although Black students were lower than White students in school and home-life self-esteem, Blacks slightly exceeded Whites in peer-group self-esteem. Here then, perhaps, was the source of their overall self-regard: disidentification with domains in which their evaluative prospects were poor (in this case, school and home life) and identification with domains in which their prospects were better (i.e., their peers).

A recent study suggests that this may be a not uncommon phenomenon. Analyzing data available from the National Educational Longitudinal Survey (National Center for Educational Statistics, 1992; a nationally representative longitudinal survey begun in 1988), Osborne (1994) found that from the 8th through 10th grades, Black students had lower achievement and somewhat higher self-esteem than White students, which replicated the general pattern of findings described above. But more than this, he found evidence of increasing Black students' disidentification over this period: The correlation between their school achievement and self-esteem for this period decreased significantly more for Black than for White students. Also, using a scale measure of school disidentification, Major, Spencer, Schmader, Wolfe, and Crocker (in press) found that Black students were more disidentified than White students in several college samples and that for disidentified students of both races, negative feedback about an intellectual task had less effect on their self-esteem than it did for identified students. Major et al. further showed that when racial stereotypes were primed, neither negative nor positive feedback affected Black students' self-esteem, whereas the self-esteem of White students followed the direction of the feedback. Ability stigmatization of the sort experienced by African Americans, then, can be associated with a protective "disconnect" between performance and self-regard, a disconnect of the sort that is consistent with disidentification theory.

Can stereotype threat directly cause this disconnect? To test this question, Kirsten Stoutemeyer and I varied the strength of stereotype threat that female test takers (Stanford students) were under by varying whether societal differences between women and men in math performance were attributed to small but stable differences in innate ability (suggesting an inherent, gender-based limit in math ability) or to social causes such as sex-role prescriptions and discrimination (suggesting no inherent, gender-based limit in math ability). We then measured their identification with math and math-related careers, either before or after they took a difficult math test. Regardless of when identification was measured, women under stronger stereotype threat disidentified with math and math-related careers more than women under

weaker stereotype threat. Although domain identification has several determinants, these findings suggest that stereotype threat is an important one of them.

### "Wise" Schooling: Practice and Policy

As a different diagnosis, the present analysis comes to a different prescription: The schooling of stereotype-threatened groups may be improved through situational changes (analogous to those manipulated in our experiments) that reduce the stereotype threat these students might otherwise be under. As noted, psychological diagnoses have more typically ascribed the problems of these students to internal processes ranging from genes to internalized stereotypes. On the face of it, at least, internal states are more difficult to modify than situational factors. Thus, the hope of the present analysis, encouraged by our research, is that these problems might be more tractable through the situational design of schooling, in particular, design that secures these students in the belief that they will not be held under the suspicion of negative stereotypes about their group. Schooling that does this, I have called *wise*, a term borrowed from Irving Goffman (1963), who borrowed it from gay men and lesbians of the 1950s. They used it to designate heterosexuals who understood their full humanity despite the stigma attached to their sexual orientation: family and friends, usually, who knew the person beneath the stigma. So it must be, I argue, for the effective schooling of stereotype-threatened groups.

Although "wisdom" may be necessary for the effective schooling of such students, it may not always be sufficient. The chief distinction made in this analysis (between those of these groups who are identified with the relevant school domain and those who are not) raises a caution. As noted, stereotype threat is not keenly felt by those who identify little with the stereotype-threatening domain. Thus, although reducing this threat in the domain may be necessary to encourage their identification, it may not be sufficient to build an identification that is not there. For this to occur, more far-reaching strategies that develop the building blocks of domain identification may be required: better skills, greater domain self-efficacy, feelings of social and cultural comfort in the domain, a lack of social pressure to disidentify, and so on.

But for the identified of these groups, who are quite numerous on college campuses, the news may be better than is typically appreciated. For these students, feasible changes in the conditions of schooling that make threatening stereotypes less applicable to their behavior (i.e., wisdom) may be enough. They are already identified with the relevant domain, they have skills and confidence in the domain, and they have survived other barriers to identification. Their remaining problem is stereotype threat. Reducing that problem, then, may be enough to bring their performance on par with that of nonstereotyped persons in the domain.

This distinction raises an important and often overlooked issue in the design of schooling for stereotype-threatened students, that of *triage*, the issue of rendering onto the right students the right intervention. Mistakes can easily be made. For example, applying a strategy to school-identified students (on the basis of their membership in a stereotype-threatened group) that assumes weak identification,



poor skills, and little confidence could backfire. It could increase stereotype threat and underperformance by signaling that their abilities are held under suspicion because of their group membership. But the opposite mistake could be made by applying a strategy that assumes strong identification, skills, and confidence to those who are actually unidentified with the relevant domain. Merely reducing stereotype threat may not accomplish much when the more primary need of these students is to gain the interests, resources, skills, confidences, and values that are needed to identify with the domain.

Some wise strategies, then, may work for both identified and unidentified students from these groups, but others may have to be appropriately targeted to be effective. I offer some examples of both types.

For both domain-identified and domain-unidentified students:

1. Optimistic teacher-student relationships. The prevailing stereotypes make it plausible for ability-stigmatized students to worry that people in their schooling environment will doubt their abilities. Thus, one wise strategy, seemingly suitable for all students, is to discredit this assumption through the authority of potential-affirming adult relationships. The Comer (1988) Schools Project has used this strategy with great success at the elementary school level, and Johnides, von Hippel, Lerner, and Nagda (1992) have used it in designing a mentoring program for incoming minority and other students at the University of Michigan. In analogous laboratory experiments, Geoffrey Cohen, Lee Ross, and I (Cohen, Steele, & Ross, 1997) found that critical feedback to African American students was strongly motivating when it was coupled with optimism about their potential.
2. Challenge over remediation. Giving challenging work to students conveys respect for their potential and thus shows them that they are not regarded through the lens of an ability-demeaning stereotype. Urie Treisman (1985) used this strategy explicitly in designing his successful group-study workshops in math for college-aged women and minorities. Taking students where they are skillwise, all students can be given challenging work at a challenging, not overwhelming, pace, especially in the context of supportive adult-student relationships. In contrast, remedial work reinforces in these students the possibility that they are being viewed stereotypically. And this, by increasing stereotype threat in the domain, can undermine their performance.
3. Stressing the expandability of intelligence. The threat of negative-ability stereotypes is that one could confirm or be seen as having a fixed limitation inherent to one's group. To the extent that schooling can stress what Carol Dweck (1986) called the *incremental* nature of human intelligence—its expandability in response to experience and training—it should help to deflect this meanest implication of the stereotype. Aronson (1996) recently found, for example, that having African American college students repeatedly advocate the expandability of intelligence to their elementary school tutees significantly improved their own grades.

For domain-identified students:



1. Affirming domain belongingness. Negative-ability stereotypes raise the threat that one does not belong in the domain. They cast doubt on the extent of one's abilities, on how well one will be accepted, on one's social compatibility with the domain, and so on. Thus, for students whose primary barrier to school identification is stereotype threat, direct affirmation of their belongingness in the domain may be effective. But it is important to base this affirmation on the students' intellectual potential. Affirming social belonging alone, for those under the threat of an ability stereotype, could be taken as begging the question.
2. Valuing multiple perspectives. This refers to strategies that explicitly value a variety of approaches to both academic substance and the larger academic culture in which that substance is considered. Making such a value public tells stereotype-threatened students that this is an environment in which the stereotype is less likely to be used.
3. Role models. People from the stereotype-threatened group who have been successful in the domain carry the message that stereotype threat is not an insurmountable barrier there.

For domain-unidentified students:

1. Nonjudgmental responsiveness. Research by Lepper, Woolverton, Mumme, and Gurtner (1993) has identified a distinct strategy that expert tutors use with especially poor students: little direct praise, Socratic direction of students' work, and minimal attention to right and wrong answers. For students weakly identified with the domain, who are threatened by a poor reputation and who probably hold internalized doubts about their ability, this Socratic strategy has the wisdom of securing a safe teacher-student relationship in which there is little cost of failure and the gradual building of domain efficacy from small gains.
2. Building self-efficacy. Based on Bandura's (1977, 1986) theory of self-efficacy, this strategy attempts to build the student's sense of competence and self-efficacy in the schooling domain. Howard and Hammond (1985) have developed a powerful implementation of this strategy for African American and other minority students, especially in inner-city public schools.

### Existence Proof: A Wise Schooling Intervention

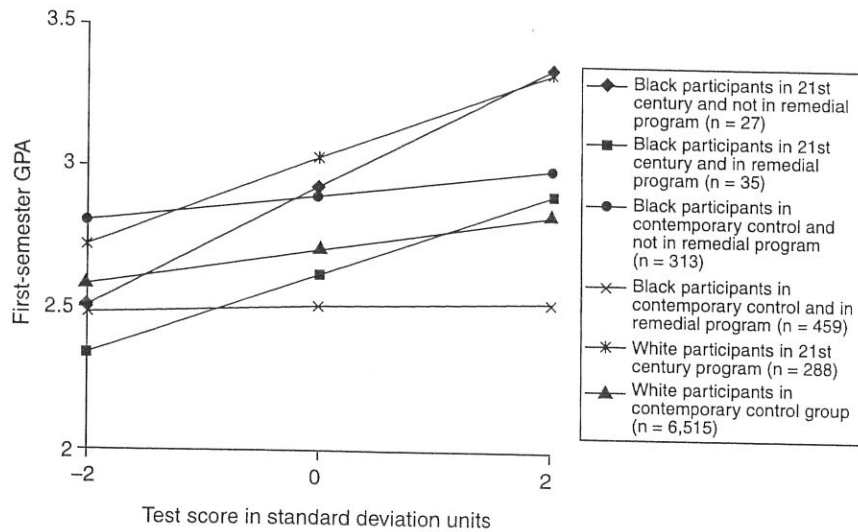
Providing a definitive test of wise schooling theory will require, of course, an extensive research program. But as a first step, something might be learned from what Urie Treisman (1985) called an existence proof, in this case, a demonstration that an intervention derived from the theory could stop or reverse a tenacious negative trajectory in the school performance of stereotype-threatened students. Such an intervention would of necessity confound things: different wise practices as well as other practices and structures, peculiar to that setting, that could also affect academic outcomes. It could not stand as a test of the psychological theory *per se*. But if a

particular architecture of wise strategies succeeded, it would encourage their applicability to the real-world schooling of these students.

With this rationale, my colleagues and I (Steven Spencer, Richard Nisbett, Mary Hummel, David Schoem, Kent Harber, Ken Carter) implemented a freshman-year program at the University of Michigan aimed at the underachievement and low retention rates of African American students. Each year, the program included approximately 250 freshmen in the ethnic proportions of the larger campus but with an oversampling of approximately 20% Black students and 20% non-Black minority students (i.e., Asian, Hispanic, and Native American students as a single group). Program students were randomly selected from the students admitted to Michigan and then recruited by phone to participate. All program participants lived together in the wing of a large, 1,200-student dormitory throughout their freshman year.

In this context, we implemented several wise strategies. The program was presented as a transition program aimed at helping students maximize the advantages of university life. We also recruited students honorifically; they were told that, as Michigan admittees, they had survived a very competitive selection process and that our program was designed to help them maximize their strong potential. These practices represented the program as nonremedial and represented the university as having acknowledged their intellectual potential and as having high expectations for them—all things that signal the irrelevance of negative group stereotypes. Once the students were in the program, these expectations were reinforced by their being offered a “challenge” workshop, modeled on those developed by Treisman (1985) for calculus, in either freshman calculus, chemistry, physics, or writing. These were taken on a voluntary basis in the dormitory. Students also participated in small weekly discussion groups, centered on brief readings, that allowed discussion of adjustment-relevant social and even personal issues. This activity has the wisdom of letting students know that they, or other members of their group, are not the only ones with concerns about adjusting to university life—an insight that can deflect the relevance of negative group stereotypes. These formal program components lasted for the first 10 weeks of the school year, and, as voluntary activities, approximately half of the students regularly participated in either one or both of them.

The first-semester grades averaged over the first two years of this ongoing project give a reliable picture of the program's initial impact. To show the size of the program's effect on students at different levels of preparation, Figure 11.5 graphs first-semester grades, using regression lines, for the different student groups as a function of standardized test scores on entry into the university (they are presented as standard deviation units in this figure to provide a common scale for students who took either the SAT or American College Test exam). The first thing to notice is the two essentially parallel lines for White and Black students outside of any program at Michigan. They replicate the standard overprediction-underperformance of Black students alluded to earlier, and it is against this pattern that the effects of the program can be evaluated. Looking first at the line for White students in our program, there is a modest tendency for these students to do better than the White control students (i.e., those outside the program), but given our accumulation of  $n$  throughout these first two years, this difference is not significant. It is the results for Black students in our program (but who were not also in the campus minority program) that are most



**Figure 11.5** First-Semester Grade Point Average (GPA) as a Function of Program and Race Controlling for High School GPA.

promising. Their line is considerably above that for Black control students (i.e., Black students outside any program) and, even with the modest sample size ( $n = 27$ ), is significantly higher than this control line in the top one third of the standardized test distribution,  $t = 2.72$ ,  $p < .05$ . It is important that this group of Black students showed almost no underperformance; in the top two thirds of the test distribution, they had essentially the same grades as White students. We also know from follow-up data that their higher grade performance continued at least through their sophomore year and that as long as four years later, only one of them had dropped out.

Theoretically just as important, is the bottom line in Figure 11.5, depicting the results for Black students in a large minority remediation program. Despite getting considerable attention, they performed worse than the other groups at nearly every level of preparation. The difference between Black students in the minority program and Black students not in any program becomes significant at 1.76 standard deviations below the mean for test performance and is significant from that point on,  $ps < .05$ . Also, by the beginning of their junior year, 25% of these students had failed to register, and among those who entered with test scores in the top one third of the test distribution, this figure was 40%. Some selection factor possibly contributed to this. Despite our having controlled for test scores and high school grade point averages in these analyses, some portion of these students may have been assigned to this program because they evidenced other risk factors. Still, these results suggest that the good intentions of the minority-remediation framework for schooling African American students can backfire by, in our terms, institutionalizing the racial stereotype by which they are already threatened.

Although these findings are preliminary and we do not know that they were mediated as our theory claims, they are a step toward an existence proof; they show that wise practices can reduce Black students' underachievement in a real-school context and, as important, that unwise practices seem to worsen it.

## Conclusion

In social psychology, we know that as observers looking at a person or group, we tend to stress internal, dispositional causes of their behavior, whereas when we take the perspective of the actor, now facing the circumstances they face, we stress more situational causes (e.g., E. E. Jones & Nisbett, 1972; Ross, 1977). If there is a system to the present research, it is that of taking the actor's perspective in trying to understand the intellectual performance of African American and female students. It is this perspective that brings to light the broadly encompassing condition of having these groups' identification with domains of schooling threatened by societal stereotypes. This is a threat that in the short run can depress their intellectual performance and, over the long run, undermine the identity itself, a predicament of serious consequence. But it is a predicament—something in the interaction between a group's social identity and its social psychological context, rather than something essential to the group itself. Predicaments can be treated, intervened on, and it is in this respect that I hope the perspective taken in this analysis and the early evidence offer encouragement.

## Notes

1. Other factors may also contribute. For example, there are persistent reports of women and minorities being treated differently in the classroom and in other aspects of schooling (e.g., Hewitt & Seymour, 1991). This treatment includes both the "chilly-climate" sins of omission—the failure to call on them in class or to recognize and encourage their talents, and so on—and, in the case of low-income minorities, sins of commission—disproportionate expulsion from school, assignment to special education classes, and administration or corporal punishment ("National Coalition of Advocates for Students Report," 1988).
2. The point is not that negative stereotypes are never internalized as low self-expectancies and self-doubts. It is that in such internalization, disidentification is the more primary adaptation. That is, once the stereotype-relevant domain (e.g., math) is dropped as a self-definition, the negative stereotype (e.g., that women are limited in math) can be accepted as more self-descriptive (i.e., internalized) without it much affecting one's self-regard (as for the woman who, not caring about math, says she is lousy at it). But this internalization is probably resisted (e.g., Crocker & Major, 1989) until disidentification makes it less self-threatening. Once this has happened, the person is likely to avoid the domain because of both disinterest and low confidence regardless of whether stereotype threat is present.
3. As a process of social devaluation, stereotype threat is both a subform of stigmatization and something more general. It is that form of stigmatization that is mediated by collectively held, devaluing group stereotypes. This means that it does not include stigmatization that derives from nonstereotyped features such as a facial disfigurement or, for example, what Goffman (1963) called abominations of the body. Stereotype threat is a situational predicament. And, in this sense, it is also more general than stigmatization. It is a threat that can befall anyone about whom a negative reputation or group stereotype exists.
4. Moreover, a protective avoidance of identification can become a group norm. In reaction to a shared sense of threat in school, for example, it can become a shared reaction that is transmitted to group members as the normative relation to school. Both research (e.g., Ogbu, 1986; Solomon, 1992) and the media have documented this reaction in minority students from inner-city high schools to Harvard University's campus. Thus, disidentification can be sustained by normative pressure from the in-group as well as by stereotype threat in the setting.
5. Participants did not actually take the test in this experiment, as completing these measures would likely have activated the stereotype in all conditions.
6. Those who are less domain identified in the stereotype-threatened group may also underperform on standardized tests. Because they care less about the domain it represents, they may be undermotivated or they may withdraw effort in the face of frustration. And for all of the reasons I have discussed, the greater portion of the stereotype-threatened group may be academically unidentified. This fact too, then, may contribute to the group's overall weaker performance on these tests in comparison with nonstereotype-threatened groups.



## References

- Adelman, C. (1991). *Women at thirty-something: Paradoxes of attainment*. Washington, DC: U.S. Department of Education, Office of Research and Development.
- Alexander, K. L., & Entwistle, D. R. (1988). Achievement in the first two years of school: Patterns and processes. *Monographs of the Society for Research in Child Development*, 53(2).
- Allport, G. (1954). *The nature of prejudice*. New York: Doubleday.
- American Council on Education. (1995-1996). *Minorities in higher education*. Washington, DC: Office of Minority Concerns.
- Aronson, J. (1996). *Advocating the malleability of intelligence as an intervention to increase college grade performance*. Unpublished manuscript, University of Texas.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1986). *Social foundations of action: A social-cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Baumeister, R. F., & Showers, C. J. (1984). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, 16, 361-383.
- Benbow, C. P., & Arjmand, O. (1990). Predictions of high academic achievement in mathematics and science by mathematically talented students: A longitudinal study. *Journal of Educational Psychology*, 82, 430-441.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210, 1262-1264.
- Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, 222, 1029-1031.
- Bettelheim, B. (1943). Individual and mass behavior in extreme situations. *Journal of Abnormal and Social Psychology*, 38, 417-452.
- Brill, A. A. (Ed.). (1938). *The basic writings of Sigmund Freud*. New York: Random House.
- Burton, N. W., & Jones, L. V. (1982). Recent trends in achievement levels of Black and White youth. *Educational Researcher*, 11, 10-17.
- Carter, S. (1991). *Reflections of an affirmative action baby*. New York: Basic Books.
- Clark, K. B. (1965). *Dark ghetto: Dilemmas of social power*. New York: Harper & Row.
- Clark, K. B., & Clark, M. K. (1939). The development of consciousness of self and the emergence of racial identification of Negro school children. *Journal of Social Psychology*, 10, 591-599.
- Cohen, G., Steele, C. M., & Ross, L. (1997). *Giving feedback across the racial divide: Overcoming the effects of stereotypes*. Unpublished manuscript, Stanford University.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Comer, J. (1988, November). Educating poor minority children. *Scientific American*, 259, 42.
- Cooley, C. H. (1956). *Human nature and the social order*. New York: Free Press.
- Cose, E. (1993). *The rage of a privileged class*. New York: Harper Collins.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608-630.
- Cross, W. E., Jr. (1991). *Shades of black: Diversity in African-American identity*. Philadelphia: Temple University Press.
- Culotta, E., & Gibbons, A. (Eds.). (1992, November 13). Minorities in science [Special section]. *Science*, 258, 1176-1232.
- DeBoer, G. (1984). A study of gender effects in science and mathematics course-taking behavior among students who graduated from college in the late 1970's. *Journal of Research in Science Teaching*, 21, 95-103.
- Demo, D. H., & Parker, K. D. (1987). Academic achievement and self-esteem among Black and White college students. *Journal of Social Psychology*, 4, 345-355.
- Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.
- Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, 11, 135-172.
- Eccles-Parsons, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectations, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). New York: Freeman.
- Erikson, E. (1956). The problem of ego-identity. *Journal of the American Analytical Association*, 4, 56-121.
- Fanon, F. (1967). *Black skins, white masks*. New York: Grove Press. (Original work published 1952)
- Frable, D., Blackstone, T., & Sherbaum, C. (1990). Marginal and mindful: Deviants in social interaction. *Journal of Personality and Social Behavior*, 59, 140-149.
- Geen, R. G. (1991). Social motivation. *Annual Review of Psychology*, 42, 377-399.
- Gerard, H. (1983). School desegregation: The social science role. *American Psychologist*, 38, 869-878.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York: Touchstone.

- Grier, W. H., & Coobs, P. M. (1968). *Black rage*. New York: Basic Books.
- Hare, B. R., & Costenell, L. A. (1985). No place to run, no place to hide: Comparative status and future prospects of Black boys. In M. B. Spencer, G. K. Brookins, & W. Allen (Eds.), *Beginnings: The social and affective development of Black children* (pp. 201-214). Hillsdale, NJ: Erlbaum.
- Herrnstein, R. A., & Murray, C. (1994). *The bell curve*. New York: Grove Press.
- Hewitt, N. M., & Seymour, E. (1991). *Factors contributing to high attrition rates among science and engineering undergraduate majors*. Unpublished report to the Alfred P. Sloan Foundation.
- Howard, J., & Hammond, R. (1985, September 9). Rumors of inferiority. *New Republic*, 72, 18-23.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139-155.
- James, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover. (Original work published 1890)
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, 1-123.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Johnides, J., von Hippel, W., Lerner, J. S., & Nagda, B. (1992, August). *Evaluation of minority retention programs: The undergraduate research opportunities program at the University of Michigan*. Paper presented at the 100th Annual Convention of the American Psychological Association, Washington, DC.
- Jones, E. E., Farina, A., Hastorf, A. H., Markus, H., Miller, O. T., & Scott, R. A. (1984). *Social stigma: The psychology of marked relationships*. New York: Freeman.
- Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning Press.
- Jones, J. M. (1997). *Prejudice and racism* (2nd ed.). New York: McGraw-Hill.
- Kamin, L. (1974). *The science and politics of I.Q.* Hillsdale, NJ: Erlbaum.
- Kardiner, A., & Ovesey, L. (1951). *The mark of oppression: Explorations in the personality of the American Negro*. New York: Norton.
- Kleck, R. E., & Strenta, A. (1980). Perceptions of the impact of negatively valued physical characteristics on social interactions. *Journal of Personality and Social Psychology*, 39, 861-873.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J.-L. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75-104). Hillsdale, NJ: Erlbaum.
- Levin, J., & Wyckoff, J. (1988). Effective advising: Identifying students most likely to persist and succeed in engineering. *Engineering Education*, 78, 178-182.
- Lewin, K. (1941). *Resolving social conflict*. New York: Harper & Row.
- Linn, M. C. (1994). The tyranny of the mean: Gender and expectations. *Notices of the American Mathematical Society*, 41, 766-769.
- Lord, C. G., & Saenz, D. S. (1985). Memory deficits and memory surfeits: Differential cognitive consequences of tokenism for tokens and observers. *Journal of Personality and Social Psychology*, 49, 918-926.
- Lovely, R. (1987, February). *Selection of undergraduate majors by high ability students: Sex difference and attrition of science majors*. Paper presented at the annual meeting of the Association for the Study of Higher Education, San Diego, CA.
- Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (in press). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Miller, L. S. (1995). *An American imperative: Accelerating minority educational advancement*. New Haven, CT: Yale University Press.
- Miller, L. S. (1996, March). *Promoting high academic achievement among non-Asian minorities*. Paper presented at the Princeton University Conference on Higher Education, Princeton, NJ.
- National Center for Educational Statistics. (1992). *National Educational Longitudinal Study of 1988: First follow-up. Student component data file user's manual*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- National Coalition of Advocates for Students Report. (1988, December 12). *The Ann Arbor News*, pp. A1, A4.
- Nettles, M. T. (1988). *Toward undergraduate student equality in American higher education*. New York: Greenwood.
- Ogbu, J. (1986). The consequences of the American caste system. In U. Neisser (Ed.), *The school achievement of minority children: New perspectives* (pp. 19-56). Hillsdale, NJ: Erlbaum.
- Osborne, J. (1994). Academics, self-esteem, and race: A look at the underlying assumption of the disidentification hypothesis. *Personality and Social Psychology Bulletin*, 21, 449-455.
- Porter, J. R., & Washington, R. E. (1979). Black identity and self-esteem: A review of the studies of Black self-concept, 1968-1978. *Annual Review of Sociology*, 5, 53-74.
- Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex,*

- language, and ethnic groups (College Board Report No. 93-1, ETS No. 94.27). New York: College Entrance Examination Board.
- Rosenberg, M. (1979). *Conceiving self*. New York: Basic Books.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 337-384). New York: Academic Press.
- Sarason, I. G. (1980). Introduction to the study of test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 57-78). Hillsdale, NJ: Erlbaum.
- Sartre, J. P. (1965). *Anti-Semite and Jew*. New York: Schocken Books. (Original work published 1946)
- Schlenker, B. R., & Weigold, M. F. (1989). Goals and the self-identification process: Constructing desired identities. In L. A. Pervin (Ed.), *Goals concepts in personality and social psychology* (pp. 243-290). Hillsdale, NJ: Erlbaum.
- Simmons, R. G., Brown, L., Bush, D. M., & Blyth, D. A. (1978). Self-esteem and achievement of Black and White adolescents. *Social Problems*, 26, 86-96.
- Solomon, R. P. (1992). *Forging a separatist culture*. Albany: State University of New York Press.
- Spencer, S., Steele, C. M., & Quinn, D. (1997). *Under suspicion of inability: Stereotype threat and women's math performance*. Manuscript submitted for publication.
- Steele, C. M. (1975). Name-calling and compliance. *Journal of Personality and Social Psychology*, 31, 361-369.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302). New York: Academic Press.
- Steele, C. M. (1992, April). Race and the schooling of Black Americans. *The Atlantic Monthly*, pp. 68-78.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811.
- Steele, S. (1990). *The content of our character*. New York: St. Martin's Press.
- Strenta, A. C., Elliott, R., Adair, R., Scott, J., & Matier, M. (1993). *Choosing and leaving science in highly selective institutions*. Unpublished report to the Alfred P. Sloan Foundation.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). New York: Academic Press.
- Treisman, U. (1985). *A study of mathematics performance of Black students at the University of California, Berkeley*. Unpublished report.
- Vars, F. E., & Bowen, W. G. (1997). *SAT scores, race, and academic performance: New evidence from academically successful colleges*. Unpublished manuscript.
- Ware, N. C., & Dill, D. (1986, March). *Persistence in science among mathematically able male and female college students with pre-college plans for a scientific major*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Ware, N. C., Steckler, N. A., & Leserman, J. (1985). Undergraduate women: Who chooses a science major? *Journal of Higher Education*, 56, 73-84.
- Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin*, 76, 92-104.
- Wylie, R. (1979). *The self-concept* (Vol. 2). Lincoln: University of Nebraska Press.